Annu. Rev. Biochem. 1989. 58:607–33

THE BACTERIAL PHOTOSYNTHETIC REACTION CENTER AS A MODEL FOR MEMBRANE PROTEINS[1]

D. C. Rees, H. Komiya, and T. O. Yeates

Department of Chemistry and Biochemistry, Molecular Biology Institute, University of California, Los Angeles, California 90024

J. P. Allen and G. Feher

Department of Physics, University of California, San Diego, La Jolla, California 92093

CONTENTS

PERSPECTIVES AND SUMMARY

Membrane proteins participate in many fundamental cellular processes. Until recently, an understanding of the function and properties of membrane pro-

[1]Abbreviations used: *Rb. sphaeroides, Rhodobacter sphaeroides; Rps. viridis, Rhodopseudomonas viridis;* RC, reaction center

0066-4154/89/0701-0607$02.00

teins was hampered by an absence of structural information at the atomic level. A landmark achievement toward understanding the structure of membrane proteins was the crystallization (1) and structure determination (2–5) of the photosynthetic reaction center (RC) from the purple bacteria *Rhodopseudomonas viridis,* followed by that of the RC from *Rhodobacter sphaeroides* (6–17). The RC is an integral membrane protein–pigment complex, which carries out the initial steps of photosynthesis (reviewed in 18). RCs from the purple bacteria *Rps. viridis* and *Rb. sphaeroides* are composed of three membrane-associated protein subunits (designated L, M, and H), and the following cofactors: four bacteriochlorophylls (Bchl or B), two bacteriopheophytins (Bphe or ϕ), two quinones, and a nonheme iron. The cofactors are organized into two symmetrical branches that are approximately related by a twofold rotation axis (2, 8). A central feature of the structural organization of the RC is the presence of 11 hydrophobic α-helixes, approximately 20–30 residues long, which are believed to represent the membrane-spanning portion of the RC (3, 9). Five membrane-spanning helixes are present in both the L and M subunits, while a single helix is in the H subunit. The folding of the L and M subunits is similar, consistent with significant sequence similarity between the two chains (19–25). The L and M subunits are approximately related by the same twofold rotation axis that relates the two cofactor branches.

RCs are the first membrane proteins to be described at atomic resolution; consequently they provide an important model for discussing the folding of membrane proteins. The structure demonstrates that α-helical structures may be adopted by integral membrane proteins, and provides confirmation of the utility of hydropathy plots in identifying nonpolar membrane-spanning regions from sequence data. An important distinction between the folding environments of water-soluble proteins and membrane proteins is the large difference in water concentration surrounding the proteins. As a result, hydrophobic interactions (26) play very different roles in stabilizing the tertiary structures of these two classes of proteins; this has important structural consequences. There is a striking difference in surface polarity of membrane and water-soluble proteins. However, the characteristic atomic packing and surface area appear quite similar.

A computational method is described for defining the position of the RC in the membrane (10). After localization of the RC structure in the membrane, surface residues in contact with the lipid bilayer were identified. As has been found for soluble globular proteins, surface residues are less well conserved in homologous membrane proteins than the buried, interior residues. Methods based on the variability of residues between homologous proteins are described (13); they are useful (*a*) in defining surface helical regions of both membrane and water-soluble proteins and (*b*) in assigning the side of these

helixes that are exposed to the solvent. A unifying view of protein structure suggests that water-soluble proteins may be considered as modified membrane proteins with covalently attached polar groups that solubilize the proteins in aqueous solution.

INTRODUCTION

Despite the central role of membrane proteins in cellular processes, comparatively little structural information has been available concerning the atomic details of their structural organization. This is in striking contrast to the situation prevailing for water-soluble proteins, for which more than 200 atomic structures have been determined. The α-helix is believed to be an important structural motif in membrane proteins, as suggested by the formation of α-helical structures by polypeptides in nonaqueous solvents (27). The first glimpse of a membrane protein structure was provided for bacteriorhodopsin from the electron microscopy studies of Henderson & Unwin (28). A striking feature of this structure was the arrangement of the membrane-spanning region of bacteriorhodopsin into seven rodlike features identified as α-helixes. Although a description of the structure at atomic resolution was not possible, this work motivated extensive efforts in understanding the folding of membrane proteins. A prominent area of this research was the development of methods for the identification and characterization of membrane-spanning helixes from sequence data (reviewed in 29, 30, 35).

This situation has changed recently with the structure determination of photosynthetic reaction centers (RCs) from purple bacteria by X-ray diffraction. RCs are integral membrane protein complexes, which carry out the initial photosynthetic steps of light-induced electron transfer from a donor to a series of acceptor species (reviewed in 18). The kinetics of these processes have been extensively studied spectroscopically, and a general picture of the structural organization of the RC was provided by a variety of biophysical, biochemical, and molecular biology studies. A critical development in obtaining detailed structural information was the crystallization (1) and structure determination (2–5) of the RC from *Rhodopseudomonas viridis* by Michel, Deisenhofer, Huber and coworkers. For the first time, X-ray crystallography revealed the atomic structure of a membrane protein, and significantly, one of great biological interest. Subsequently, the structure of the RC from *Rhodobacter sphaeroides* has been described (6–17). Knowledge of these structures has opened a window into the molecular architecture of membrane protein systems.

Since the RC structure determinations, a number of conference proceedings and reviews have appeared that discuss various aspects of the photosynthetic function of the RC (31–34). Many important studies related to electron

transfer and spectroscopic studies of RCs are described in these collections. This review focuses only on aspects of the RC structure relevant to the folding of membrane proteins. Emphasis is placed on the surface properties of the membrane-spanning region of the RC, and comparison between folding characteristics of membrane and water-soluble proteins. The discussion is based on structural and sequence information obtained for RCs from purple bacteria; these are to date the best-characterized membrane protein system.

Notation

The following conventions are used throughout this review: the three RC protein subunits are designated L, M, and H. Residue numbers are based on the *Rb. sphaeroides* protein sequences (19–21). Cofactors are abbreviated as follows: bacteriochlorophyll (Bchl or B), bacteriopheophytin (Bphe or ϕ), quinone (Q), iron (Fe), and carotenoid (C). The special pair of bacteriochlorophyll is represented as $(Bchl)_2$ or D. The two cofactor branches are designated A and B (8), which correspond to the L and M branches, respectively, in the description of the RC from *Rps. viridis* (2–4). Branch affiliations of the different cofactors are indicated by a subscript A or B. LDAO and BOG represent the detergent molecules lauryl dimethylamine-N-oxide and β-octyl glucoside, respectively. The term "soluble protein" specifically refers to globular proteins soluble in aqueous solutions in the absence of detergents or membranes.

STRUCTURE OF THE RC

RCs from purple bacteria such as *Rps. viridis* and *Rb. sphaeroides* typically contain one copy each of three membrane-associated subunits designated L, M, and H (reviewed in 18). These labels (Light, Medium, and Heavy) were assigned on the basis of apparent mobilities on SDS gels, although subsequent sequence analyses indicated that the actual molecule weights increase in the order H < L < M. Certain RCs, including the one from *Rps. viridis* (but not from *Rb. sphaeroides*), contain a tightly associated cytochrome molecule. Amino acid sequences have been determined for RCs from four purple bacteria, *Rb. sphaeroides* (19–21), *Rhodobacter capsulatus* (22), *Rps. viridis* (23, 24), and *Rhodospirillum rubrum* (25). Determination of the primary sequences for these proteins provided important structural information. Sequence homology between the L and M subunits was evident, thus suggesting that they originally were derived from a common precursor subunit. Hydropathy plots (29, 30, 35) of the amino acid sequences of the L, M, and H subunits indicate a total of 11 membrane-spanning helixes in the RC; the analysis was based on the occurrence of stretches of 20–30 predominantly hydrophobic residues. The L and M subunits each exhibit five such regions, while the H subunit contains only one region.

Associated with these proteins are a group of cofactors, including four Bchl, two Bphe, two quinones, and a nonheme iron. Two of the Bchl are organized into the special pair or dimer, $(Bchl)_2$, which is the primary electron donor for the photosynthetic reactions. The electron transfer reaction proceeds from the dimer to an intermediate acceptor (ϕ_A), a primary quinone (Q_A), and a secondary quinone (Q_B). Also associated with RCs is a single carotenoid molecule, except in certain carotenoidless mutants such as *Rb. sphaeroides* strain R-26.

Detailed structural determinations of membrane proteins, including RCs, require high-resolution crystallographic analyses, which in turn require the availability of suitable crystals. Ordered two-dimensional arrays of membrane proteins, in particular bacteriorhodopsin (28), have provided valuable structural information, but these analyses have not progressed to the approximately 3 Å resolution required for atomic model building. Garavito & Rosenbusch (36), and Michel (1) succeeded in obtaining well-ordered three-dimensional crystals of two membrane proteins: porin from *Escherichia coli* and the RC from *Rps. viridis*, respectively. These accomplishments opened the field of X-ray structural analysis of membrane proteins. Approaches to the crystallization of membrane proteins have been reviewed (37, 38, 38a).

The structure of the RC from *Rps. viridis* was initially determined to 3 Å resolution (2), using X-ray diffraction techniques and the method of multiple isomorphous replacement. The resolution and refinement of this structure has been subsequently extended to 2.3 Å (5). Crystallization of the carotenoidless strain R-26 (6, 6a, 15), wild-type strain 2.4.1 (14, 39), and strain Y (40) of the RC from *Rb. sphaeroides* have been reported. We have determined structures of the *Rb. sphaeroides* RCs from the carotenoidless strain R-26 and the wild-type strain 2.4.1 of *Rb. sphaeroides* at 2.8 Å and 3.0 Å resolution, respectively (8–14). This model of the RC from *Rb. sphaeroides* R-26 is used for the figures and analyses described in this review. A second, independent structure determination of the RC from *Rb. sphaeroides* R-26 has also been presented at 3.2 Å resolution (16, 17). The structure of the RC from *Rb. sphaeroides* is similar to that of the RC from *Rps. viridis;* this allowed the use of the method of molecular replacement to solve the initial crystallographic phase problem (7, 16).

The overall folding of the polypeptide chains and the cofactor arrangement in the *Rb. sphaeroides* R-26 RC is illustrated in Figure 1. A striking feature of the RC is the approximate symmetry evident in the structure, with both the two cofactor branches and the L and M subunits related by a twofold rotation axis. In view of the complexity of the interactions, it is convenient to describe initially the structures of the protein and cofactors separately. It must be kept in mind, however, that the cofactors and the protein subunits are intimately associated with each other, and it is unlikely that structures of any individual component could be maintained in isolation.

Protein Structure

A total of 11 hydrophobic α-helixes are observed in the L, M, and H subunits; they create a framework that organizes the cofactors (Figures 1 and 2). The general positions of the helixes in the protein sequences had been predicted from hydropathy analyses. Since these helixes create an apolar region approximately 35 Å wide, it is assumed that they represent the membrane-spanning region of the RC (2, 3, 9). The L and M subunits each have five apolar helixes designated A, B, C, D, and E, preceded by the appropriate subunit letter (i.e. LA, LB, etc). The H subunit has one membrane-spanning helix designated HA. The LD, LE, MD, and ME helixes form a core structure that interacts extensively with the cofactors. In contrast, the A, B, and C helixes of the L and M subunits, and the HA helix are located on the periphery of the RC, away from the cofactor rings. The tilts and curvatures of these helixes (3, 9) influence the number of residues necessary to span the membrane. The average tilt angle of the 11 membrane-spanning helixes is 22° from the twofold axis that relates the L and M subunits; the axes of individual helixes may differ by up to 65° (the angle between the LD and MD helixes). In addition, the C and E helixes exhibit substantial kinking, and have effective radii of curvature of 30–70 Å.

More than half of the residues in the RC, including most of the H subunit, are located outside the membrane-spanning region. The orientation of the RC in the cell membrane has been established by chemical modification, and by studies of cytochrome c binding (reviewed in 18, 43). The studies permit assignment of the "top" surface of the RC (as indicated in Figure 1) to the periplasmic side of the membrane, while the "bottom" surface faces the cytoplasm. Helical segments in these outer regions are designated in relation

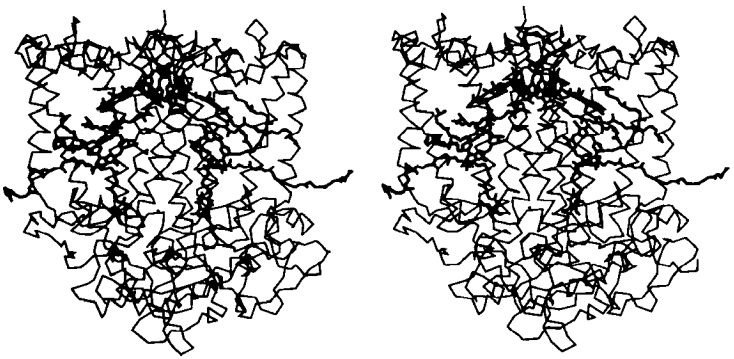

Figure 1 Stereoview of the cofactors and Cα backbone of the protein subunits of the RC from the carotenoidless mutant strain R-26 of *Rb. sphaeroides*. The twofold axis is aligned vertically in the paper, with the cytoplasmic side of the RC at the bottom of the figure. The figure was prepared with the FRODO graphics program (41). Modified from Ref. 9.

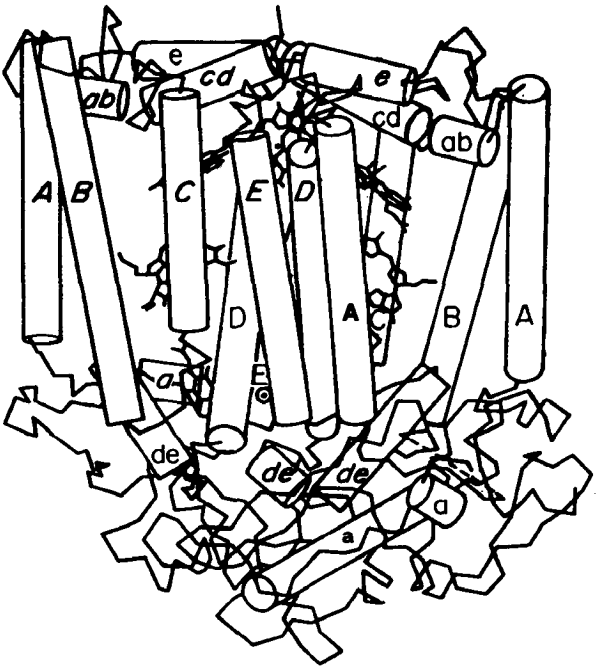

Figure 2 The RC structure with protein subunits and cofactors. The α-helixes have been approximated by straight cylinders. Helixes of the L subunit are lettered in plain type, while helixes of the M subunit are lettered in italic type. H subunit helixes (**A** and **a**) are in bold font. The phytyl and isoprenoid tails of the cofactors have been truncated. The view is in the same orientation as Figure 1. The figure was modified from Ref. 9 and was prepared with the aid of a program described in Ref. 42.

to the transmembrane helixes that they connect. An interesting arrangement on the surface of the periplasmic side of both the L and M subunits are two "interrupted" helixes [designated as the I helixes (9)] composed of three helical segments connecting helixes A and B, C and D, and one following helix E (Figure 2). These helical segments are designated ab, cd, and e, respectively. The bulk of the H subunit lies on the cytoplasmic side, and consists of several β-sheets organized into a globular domain.

Cofactor Structure

The cofactors are arranged into two branches (2, 8), designated A and B, illustrated in Figure 3. For clarity, the phytyl and isoprenoid tails have been removed from the cofactors. Along either branch, the sequence of cofactors is $(Bchl)_2$, Bchl, Bphe, Q, and Fe. The ring centers of adjacent cofactors are separated by approximately 10–13 Å (2, 8). Individual distances between cofactor centers are listed in Table 1 of Ref. 8. The RC models have firmly

established the structural basis for two important points that had been deduced spectroscopically: (*a*) the dimer interaction of two Bchl molecules in the special pair $(Bchl)_2$, whose existence had been proposed on the basis of EPR (44, 45) and ENDOR (46, 47) experiments, and (*b*) the sequence of the electron transfer steps in the order $(Bchl)_2$ to Bphe to Q. An unexpected aspect of the RC structures, however, is the existence of two cofactor branches, although the spectroscopic evidence indicates the existence of only one photochemically active branch. The active branch has been identified as the A branch (2, 8, 48).

The cofactor rings in each branch are approximately related by the same twofold rotation axis that relates the L and M subunits (3, 9). The $(Bchl)_2$ and the Fe are positioned close to this axis. Atoms in the cofactor rings can be superimposed to within 0.8 Å–1.5 Å by a twofold rotation of one branch about this axis. Significant departures from this symmetrical arrangement are observed for the phytyl and isoprenoid tails.

In addition to the cofactors associated with the primary photochemical events, the RC binds two other classes of molecules: carotenoid and detergents.

1. carotenoid: RCs bind a single carotenoid molecule, except in the carotenoidless mutant strain R-26 of *Rb. sphaeroides*. Carotenoids are long, conjugated polyenes that have been located in the structures of the RCs from both *Rps. viridis* (5) and *Rb. sphaeroides* (11). The carotenoid spher-

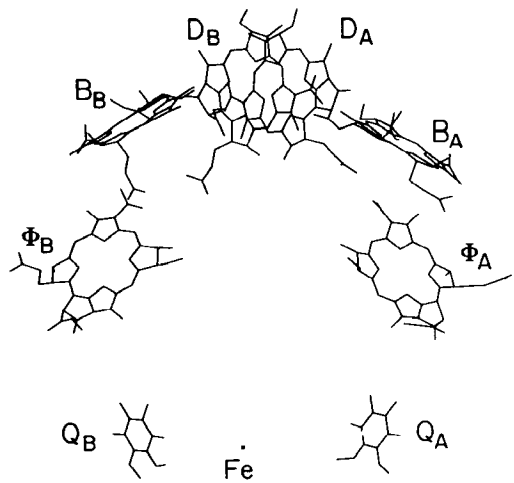

Figure 3 Cofactor structure of the RC from *Rb. sphaeroides* R-26. Phytyl and isoprenoid tails of the cofactors have been omitted for clarity. The cofactors are displayed in the same orientation as shown in Figure 1.

oidene in the RC from *Rb. sphaeroides* strain 2.4.1 adopts a boomerang-shaped structure that curves around the MA and MB helixes, passing near the B_B ring.

2. detergent: Most of the detergent molecules present in the crystal lattice are too disordered to observe crystallographically. A few exceptions have been noted, however. In the RC from *Rps. viridis,* a molecule of LDAO has been identified between the MD and HA helixes (J. Deisenhofer, personal communication). Two molecules of BOG have been identified in the structure of the RC from the carotenoidless strain R-26 of *Rb. sphaeroides* (11). One BOG molecule binds to the site occupied by the carotenoid in the RC from *Rb. sphaeroides* 2.4.1. This provides an illustration of how detergent molecules may influence the properties of membrane proteins by occupying specific ligand-binding sites. A second, less well defined, BOG is bound to the L subunit, near B_A and ϕ_A. Spectral changes observed in RCs solubilized in different detergents (49) may be due to direct interaction between detergent molecules and cofactors.

RC AS A MODEL FOR THE FOLDING OF MEMBRANE PROTEINS

With the exception of the RC, our knowledge of the tertiary structure of membrane proteins is primitive relative to that of water-soluble, globular proteins. A general picture of soluble proteins has emerged over the past 30 years, which emphasizes (*a*) the efficiently packed, relatively nonpolar interior, and (*b*) the polar surface, which minimizes the surface energy in an aqueous environment. These characteristics are a direct consequence of the influence of hydrophobic interactions (26). An important distinction between the folding environment of soluble proteins and membrane proteins is the relative absence of water in the bilayer region. Since this results in a much different role for hydrophobic interactions in stabilizing the structure of membrane proteins, a comparative analysis of the structures of proteins in these two classes is central to understanding the role that the solvent plays in protein folding. Although the RC may not prove to belong to the most prevalent class of membrane proteins (given the large cofactor component), it currently represents the best-defined structure and consequently commands our interest as a model to investigate the structural organization of membrane proteins.

Surface Area and Volume of the RC

A view of membrane protein structure in terms of a collection of hydrophobic, membrane transversing α-helixes was satisfyingly confirmed by the three-dimensional structures of the RC. As in the case of soluble proteins, it would, however, be surprising if there were only one structural motif that charac-

terized the folding of transmembrane proteins. Given the observed (or anticipated) diversity in the detailed structure of the polypeptide chains of soluble and membrane proteins, how can a general comparison of structures between these two classes be achieved? Clearly, a characterization of protein tertiary structures that is independent of the details of the local folding pattern of the polypeptide is required.

One approach to this problem utilizes the concepts of molecular surface area and volume introduced and implemented by Richards (50). These ideas have provided a quantitative basis for characterizing soluble proteins with efficiently packed, apolar interiors and a polar surface for favorable solvent interactions. The RC structure provides an opportunity to perform comparable analyses on a membrane protein.

SURFACE AREA The energy required to increase the surface area of a liquid is given by the product of the surface tension of the liquid and the change in surface area. The decreased surface tension of hydrocarbon liquids (~30 cal $Å^{-2}$) compared to water (105 cal $Å^{-2}$) (Ref. 51) suggests that an increase in surface area of the solvent surrounding a protein would require less energy in a membrane than in water. Consequently, membrane proteins might be expected to have a larger surface area than soluble proteins of the same molecular weight (if, for example, more favorable packing contacts could be achieved at the expense of a higher surface energy). Following Richards, surfaces of proteins may be defined by rolling a spherical probe around the van der Waals surface of a protein. The accessible surface area, A_s, is determined from the area of the surface generated by the center of the probe (52). For this calculation, van der Waals radii for the protein atoms were taken from Richmond & Richards (53), while all cofactor atoms were assigned radii of 1.8 Å. With a solvent probe radius of 1.4 Å, A_s for the *Rb. sphaeroides* RC was calculated to be 34,800 $Å^2$. The corresponding value for soluble proteins was estimated by an empirical relationship between A_s and molecular weight M established by Miller et al (54) for soluble, oligomeric proteins:

$$A_s = 5.3\, M^{0.76} \qquad 1.$$

For a protein with $M = 10^5$ (the molecular weight of the RC from *Rb. sphaeroides*), Equation 1 predicts A_s = 33,400 $Å^2$. The agreement between the calculated surface area (based on observations on soluble proteins) and the observed value for RC indicates that there is no significant difference in surface area between membrane and soluble proteins of comparable size. The surface energies of soluble and membrane proteins must be similar, despite the differences in surface tensions between hydrocarbon liquids and water.

Surface energies are influenced by both the solvent surface tension, and the protein-solvent interaction energy (51). Comparable surface areas for membrane and soluble proteins could reflect weaker (van der Waals) interactions between proteins and hydrocarbons in the membrane, compared to stronger (hydrogen bond) interactions possible between proteins and water in aqueous solutions.

Unlike for a smooth sphere, the surface areas of objects with irregular or rough surfaces (including macromolecules) are not uniquely defined. Hence, it is essential that surface area comparisons of proteins be performed with identical van der Waals and probe radii. A measure of the roughness of protein surfaces can be derived from the dependence of the surface area on the parameters of the calculation. The accessible surface area is unsuitable for this analysis, however, and it is necessary to adopt a second type of surface, the molecular surface. As described by Richards (50), the molecular surface is defined as a continuous envelope stretched over the van der Waals surface of a protein; it describes the position of the inner surface of the probe sphere as the probe moves in contact with the van der Waals surface. In contrast, the accessible surface is defined by the position of the center of the probe sphere as the probe moves in contact with the van der Waals surface; hence the accessible surface is always displaced from the van der Waals surface of the protein. Since the displacement varies with the probe radius, the value of the calculated surface area will be affected by both the surface roughness and the displacement. The variation in molecular surface area with probe radius, however, directly provides information on surface roughness. For a perfectly smooth object such as a sphere, the molecular surface will be independent of the probe radius. In contrast, the molecular surface area of an irregular object, such as a sponge, will depend significantly on the probe radius. A larger surface area will be calculated for small probes that can penetrate into the pores of the sponge, as compared to larger probes which are excluded from the pores.

The surface roughness of a protein may be characterized by a parameter, D, which is calculated from the variation in molecular surface area, A, with probe radius, r, through the relationship (55, 56):

$$D = 2 - \mathrm{d}(\log A)/\mathrm{d}(\log r) \qquad 2.$$

As the surface becomes more irregular, D increases from the value 2 for a smooth surface, to a value $D \leq 3$. D has properties of a (noninteger) dimension known as a "fractal dimension," which has found application in the description of a wide range of physical and mathematical phenomena (55). D is defined only for a certain range of probe radii; in the limit of both small and large probe radii, D will equal 2 for macromolecular surfaces. For small probe radii, the probe interacts predominantly with the spherical van der Waals

spheres describing the protein atoms, whereas large probes are sensitive only to the overall shape of the molecule. D will be maximal for probe radii in the size range of the irregular surface features. For a protein, this corresponds to the approximate size of water molecules and side chains, with 1.5 Å $< r <$ 3 Å.

The variation in A with r is illustrated in Figure 4 for a representative sample of both monomeric and oligomeric soluble proteins, and the RC from *Rb. sphaeroides*. Curves are presented for both the entire RC molecule, as well as atoms located in the membrane-spanning region (defined more pre-

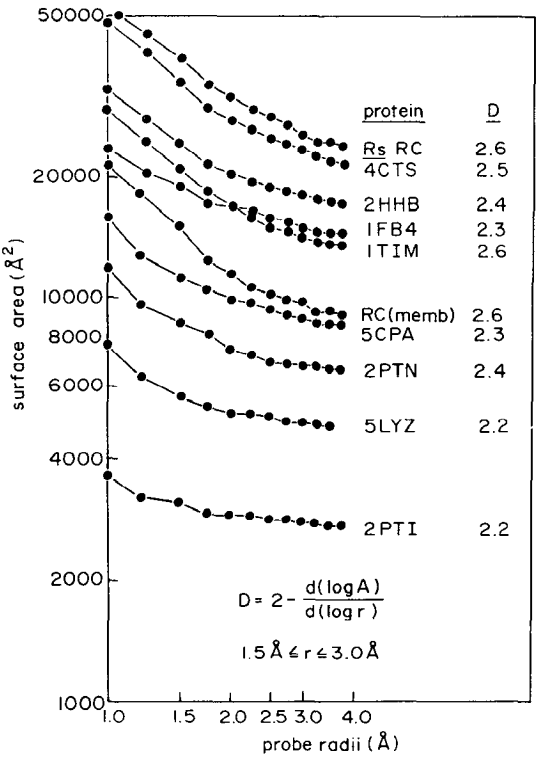

Figure 4 Variation in molecular surface area with probe radii plotted on a log-log scale for selected monomeric and oligomeric proteins. Molecular surface areas were calculated with Connelly's MS program (57). The D value (Eq. 2) provides a measure of the overall surface roughness of each protein. D was calculated with probe radii in the range 1.5 Å $< r <$ 3 Å. The following coordinate sets for soluble proteins were used from the Brookhaven Protein Data Bank (58): 2PTI (pancreatic trypsin inhibitor), 5LYZ (lysozyme), 2PTN (trypsin), 5CPA (carboxypeptidase A), 1TIM (triose phosphate isomerase), 1FB4 (immunoglobulin Kol), 2HHB (hemoglobin), 4CTS (citrate synthase). Surface areas for the RC from *Rb. sphaeroides* were calculated for both the entire RC (Rs RC), and for atoms in the membrane-spanning region [RC (memb)].

cisely below). Programs developed by Connelly (57) were used for the area calculations. Van der Waals radii for the protein atoms were assigned the values 1.8 Å, 1.7 Å, 1.4 Å, 1.8 Å, and 0.8 Å for carbon, nitrogen, oxygen, sulfur, and metal atoms, respectively. (Differences in these curves and those reported in Ref. 56 are due to the use of smaller van der Waals radii in the earlier work.) The trend apparent from this figure is that larger (oligomeric) proteins have larger values of D, i.e. they are more irregular than smaller proteins. Thus, large proteins are not simply smaller proteins scaled up in size. There is no significant difference in overall surface roughness of the RC relative to soluble, oligomeric proteins of comparable size. Previous studies (56) indicated that the surfaces of soluble proteins are not uniformly irregular, but rather exhibit variations in roughness between different regions of the protein surface. Comparable studies examining the local variations in surface roughness of the RC structure have not yet been performed.

ATOMIC VOLUMES AND PACKING The volumes of buried atoms in a protein may be calculated with the Voronoi construction (50, 59, 60). In this method, planes are drawn that are perpendicular bisectors to all the vectors between pairs of atoms in the structure. These planes intersect to define a unique polyhedron around each atom. Only buried atoms (with zero accessible surface area) are included in the calculation, to ensure that a closed polyhedron with a defined volume is constructed. The atomic volume is defined by the volume of the polyhedron surrounding the atom. An important conclusion from volume calculations is that the packing density of buried atoms in soluble proteins is the same as that observed in crystals of small organic molecules; i.e. interior atoms in soluble proteins are efficiently packed (50). Volumes of buried atoms in the membrane-spanning region of the RC have been calculated, and are similar to those observed for interior atoms in soluble proteins such as carboxypeptidase A (Table 1) and ribonuclease S (60). Consequently, the same efficient packing that characterizes soluble proteins is also maintained in the RC structure.

Stabilization of the Tertiary Structure of Membrane Proteins

The RC maintains a well-defined tertiary structure in the membrane-spanning region, despite the decrease in significance of hydrophobic interactions relative to soluble proteins. Based on the RC structure, the following types of interactions appear to impart the necessary structural specificity in the transmembrane region (10):

1. Atomic packing in the transmembrane region. The observed efficient packing of atoms in the RC structure stabilizes the tertiary structure by maximizing van der Waals contacts between atoms, and minimizing the adverse consequences of cavities (61).

Table 1 Volumes (Vol) with standard deviations (SD) of buried atoms in the membrane-spanning region of the RC from *Rb. sphaeroides* and the water soluble globular protein carboxypeptidase A (adapted from Ref. 10)

	RC		Carboxypeptidase A	
Atom type	Vol, Å^3	SD, Å^3	Vol, Å^3	SD, Å^3
Main-chain atoms				
N	13	2	14	2
Cα	12	3	12	2
C	8	1	8	1
O	21	4	22	3
Pro N	10	1	10	1
Side-chain atoms				
CβH	13	3	13	1
CβH$_2$	21	6	23	8
CH	21	2	21	3
CH$_2$	14	3	14	2
CH$_3$	31	6	34	5
Aromatic C	19	7	18	5
His ring	16	4	15	4
OH	25	4	24	5
O/N	21	5	24	4
Trp	16	6	17	5

2. Polar interactions between transmembrane helixes. A major polar interaction that will stabilize the transmembrane helical arrangement is provided by the four histidine ligands on the D and E helixes, which coordinate the iron atom. More general types of electrostatic effects, such as helix dipole interactions, may be involved in stabilizing the dominantly antiparallel arrangement of the transmembrane helixes (62). On average, less than one interhelical hydrogen bond is present between the polar side chains of residues on different helixes. No salt bridges between membrane helixes are observed.

3. Protein structures outside the membrane-spanning region. Several types of organized protein structures are observed in regions of the RC exposed to the aqueous environment (Figure 2). The two periplasmic I helixes on both the L and M subunits may serve as a strap that holds the transmembrane helixes together on the periplasmic sides. Structures such as the β-sheet region, as well as contacts between the L and M subunits and the H subunit, may also stabilize the membrane-spanning structure on the cytoplasmic side. However, the H subunit does not seem to be essential for maintaining the RC structure, since its removal does not significantly change the kinetics of electron transfer up to (and including) the reduction of the primary quinone (49). Furthermore, the RC structure in the green bacterium *Chloroflexus aurantiacus* is stable despite the absence of an H subunit (63).

Membrane vs Soluble Proteins: Analogy to Crystal Morphology

Water-soluble and membrane (RC) proteins seem similar in terms of the geometrical criteria of surface area and volume. The most striking difference between these two classes of proteins is the chemical nature of the exposed surface groups. To minimize surface energies, soluble proteins fold to generate a polar surface, while membrane proteins require an apolar surface. This behavior is similar to the effect of solvent conditions on the morphology of small molecule crystals. Gibbs demonstrated that the equilibrium morphology of a crystal will have the minimum surface free energy (64). Since different crystal faces have different exposed chemical groups, changing solvent conditions will alter the crystal morphology so as to maintain the state with lowest surface energy. For example, polar crystal faces with exposed carboxyl groups dominate the morphology of succinic acid crystals grown from water, whereas more apolar crystal faces with exposed methylene carbons are prominent in crystals grown from apolar solvents or by sublimation (65). The interior packing of succinic acid molecules remains, however, unchanged under these different solvent conditions. Thus, crystal morphology may be viewed as being analogous to the "morphologies" of water-soluble and membrane proteins; i.e. the surface composition of proteins is sensitive to solvent (i.e. water or bilayer) conditions, but the same type of efficient package is maintained. This behavior suggests that water-soluble proteins may be considered as modified membrane proteins with covalently attached polar groups that make the proteins soluble in aqueous solutions.

Position of the RC in the Membrane

An important aspect of the characterization of membrane proteins is to define the region of interaction between the protein and the membrane. For the RC, this region is composed of contiguous stretches of 20–30 apolar residues, which were identified from an analysis of the sequence data. The three-dimensional structures of the RC strongly support these assignments by demonstrating that these apolar regions are organized into 11 α-helixes, that create a hydrophobic band approximately 35 Å wide. This band is essentially devoid of charged residues. Satisfying as this picture is, a direct demonstration of the membrane-spanning region of the RC is, however, difficult to achieve. In both the *Rps. viridis* and *Rb. sphaeroides* crystals, the phospholipids have been replaced by detergent molecules, which, with one or two exceptions, are disordered and therefore not observable by X-ray diffraction. A general location of the disordered detergents has been obtained by low-resolution neutron diffraction studies of the RC from *Rps. viridis* (65a). These studies localized the binding region of the detergent on the surface of the RC that surrounds the 11 α-helixes. The precise location of the boundaries of the

detergent-binding region (assuming the boundary is sharply defined), and by inference, the membrane-spanning region of the RC, could not be determined at the 15 Å resolution of this neutron diffraction study.

The position of the membrane-spanning region of the RC was determined indirectly by an analysis of the energetics of the RC-membrane interaction (10). It is based on the decrease in hydrophobic free energy when nonpolar regions of a membrane protein are placed into a lipid bilayer. Various potential functions have been developed to estimate the free energy of transfer, ΔG_H, between apolar and aqueous solvents. In general, ΔG_H is expressed as a product of two terms: (*a*) the surface area of the region involved in the transfer between solvents and (*b*) a surface free energy term. Following Eisenberg & McLachlan (66), ΔG_H may be expressed as a sum involving the solvent accessible surface area of an atom i, A_{si}, and the surface free energy $\Delta\sigma_i$ for each atom type:

$$\Delta G_H = \sum_i \Delta\sigma_i A_{si} \qquad 3.$$

where the sum is over all atoms i. The surface free energies of transfer between a nonpolar and an aqueous phase for different atom types are (in cal $Å^{-2}$) $\Delta\sigma(C) = 16$; $\Delta\sigma(N, O) = -6$; $\Delta\sigma(O^-) = -24$; $\Delta\sigma(N^+) = -50$; $\Delta\sigma(S) = 21$. These values were determined empirically (66) by fitting experimentally obtained values for the free energy of transfer of amino acids to an energy function similar to Equation 3.

With Expression 3 for ΔG_H, the equilibrium position of the RC in a bilayer can be established by determining the position of minimum energy (subject to the assumptions and limitations described below). An initial estimate of the location of the membrane-spanning region of the RC was determined by evaluating ΔG_H for sections of the RC that were 5 Å thick (Figure 5). For these calculations, the RC was sectioned normal to the local twofold axis (defined as the *z* axis), with the Fe atom at the origin ($z = 0$). The values of ΔG_H provide an estimate of the free energy of transfer from the membrane to water of the surface atoms in a particular 5 Å thick section. A region of the RC approximately 40 Å thick exhibits a large hydrophobic energy ΔG_H; this presumably represents the membrane-spanning region. Integration of the area under the curve of Figure 5 (correcting for the 5 Å slab width) yields a value of about 20 kcal/mole for ΔG_H per helix. This is consistent with an estimate of 30 kcal/mole for a single transmembrane helix (30). More detailed calculations support the near coincidence of the twofold axis with the membrane normal (10). Accordingly, the membrane-spanning region of the RC from *Rb. sphaeroides* is approximately 40 Å wide, and extends from the Fe atom on one side (cytoplasmic), to a position approximately 10–15 Å beyond the center of the dimer on the opposite (periplasmic) side.

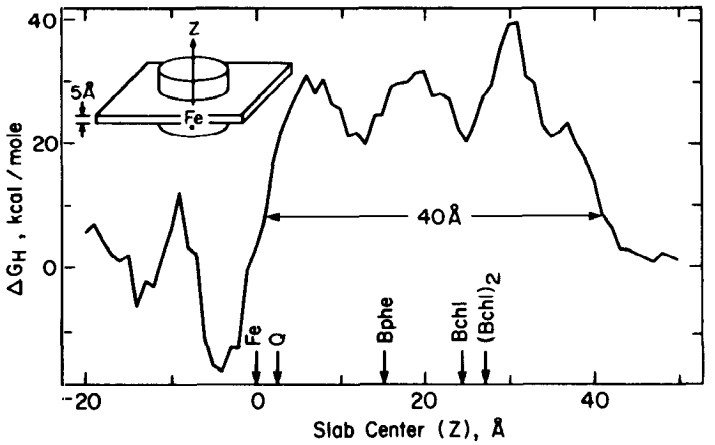

Figure 5 The energy, ΔG_H (Eq. 3), required to transfer a 5 Å thick section of the RC from the membrane to water for different positions (in 1 Å increments) of the RC. The normal of the section is parallel to the twofold symmetry axis z of the RC, as indicated schematically (inset). The projected locations of the centers of the cofactors onto the z axis are indicated. The position of the Fe was arbitrarily chosen as zero. Arrow (40 Å) indicates the membrane-spanning region. The energy calculations were performed on the experimentally determined three-dimensional structure of the RC from *Rb. sphaeroides*. Figure modified from Ref. 10.

What exactly does this 40 Å wide region represent? The Eisenberg & McLachlan expression for ΔG_H estimates the free energy of transfer from a nonaqueous to a water environment (66). Hence, the 40 Å wide slab should represent the total extent of the RC surface shielded from water. This would include both the region of the RC in contact with the fatty acid tails, as well as the region in contact with the polar head groups. It might seem surprising that the region of the RC in contact with the polar head groups would exhibit a positive ΔG_H for transfer to water, since the head groups themselves might be expected to resemble water more closely than an apolar solvent. As a relevant model, it is instructive to consider the interactions between carbohydrate and protein in glycoproteins. X-ray structures of glycoproteins (67, 68) have shown that the carbohydrate chains interact directly with parts of the protein surface, shielding those regions from exposure to water. Although carbohydrate groups are highly polar (69), they cover regions of the protein surface that have significant numbers of hydrophobic residues (67, 68). Apparently, even these polar molecules have nonpolar surface regions that interact preferentially with hydrophobic residues of the protein (69a). Consequently, a positive ΔG_H for the region of the RC in contact with the polar head groups seems plausible. Since the head group layer has a thickness of 5 Å (68), this region of the RC surface would consist of surface groups with z values of approximately 0–5 Å and 35–40 Å, with the convention illustrated in Figure

5. The region of the RC interacting with the nonpolar, fatty acid tail part of the bilayer of the RC will then extend over z = 5–35 Å.

The accuracy with which the membrane-spanning region of the RC can be identified depends critically on the assumptions made in the analysis. These assumptions fall into three general categories (10):

1. The membrane may be approximated by a planar slab with uniform thickness on all sides of the RC.

2. There is a sharp interface between the membrane and the aqueous solution, as well as between the RC and the membrane. Interactions between the RC and other proteins (such as the antennae complex) are neglected in this model. Similarly, the possible penetration of water into the bilayer region has been neglected.

3. Only hydrophobic energies are explicitly considered in this model. Other contributions to membrane–protein energetics, including electrostatic image charges arising at the protein–membrane–water boundaries, have been neglected.

A measure of the validity of these assumptions can be obtained by comparing the calculated width of the membrane with the experimentally determined values. Small angle X-ray scattering studies of *Rb. sphaeroides* vesicles yield a membrane thickness of 45 ± 5 Å (70). Likewise, small angle scattering studies of vesicles containing vaccenic acid [the major fatty acid component in the membrane of *Rb. sphaeroides* (71)] indicate that the total membrane thickness of these vesicles is 38 Å (72). The agreement between the calculated and observed values for membrane thickness indicates that the assumptions made in this analysis are reasonable approximations.

Identification of Residues Exposed to the Membrane by Sequence Analysis

In addition to the position of the transmembrane helixes in the membrane discussed in the previous section, the characterization of residues in contact with the lipid bilayer is also important for structural analyses of membrane proteins (10). Figure 6 depicts the approximate position of Cα atoms of residues of the 11 transmembrane helixes in the membrane. As described above, the nonpolar region of the bilayer extends over z = 5–35 Å, while the head groups are located in the regions z = 0–5 Å and z = 35–40 Å. Most of the charged groups found within the membrane-spanning region are contained in the head group zones.

Residues on each helix that are in contact with the membrane bilayer were identified by tabulating the accessible surface area for each residue (10). Residues with more than half of their accessible surface area exposed to the membrane are circled in Figure 6, while residues with 20–50% of their surface area exposed to the membrane are capped with a semicircular arc. The

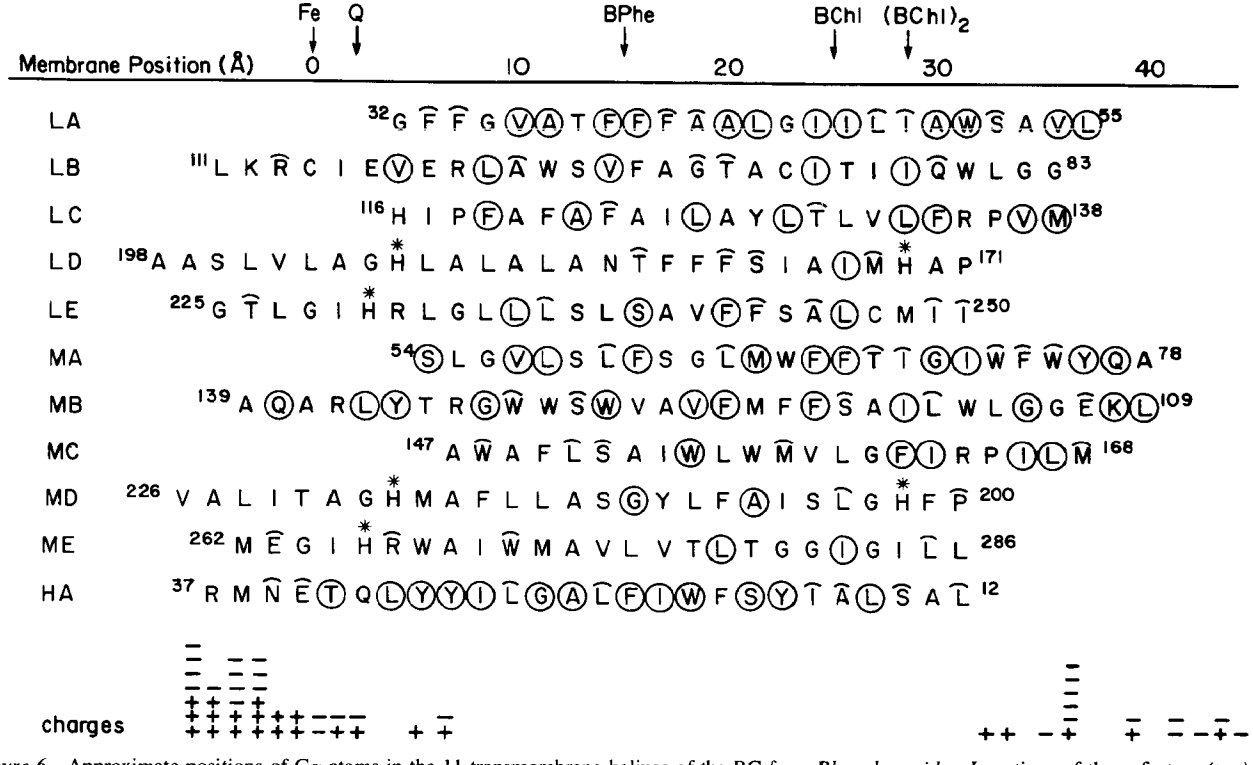

Figure 6 Approximate positions of Cα atoms in the 11 transmembrane helixes of the RC from *Rb. sphaeroides*. Locations of the cofactors *(top)* and all charged residues *(bottom)* of the RC are indicated. Residues in circles have >50% of their surface area exposed to the membrane, while capped residues have 20–50% of their surface area exposed to the membrane. The remaining residues are primarily buried inside the protein. Histidine residues marked with an asterisk are ligands to Fe or Bchl. Amino acids are designated by the single-letter code. Figure modified from Ref. 10.

remaining residues are primarily buried inside the RC. Helixes on the periphery of the RC (A, B, and C) have more residues exposed to the membrane than the core helixes (D and E). In many instances, residues exposed to the membrane are spaced at multiples of three to four residues, which corresponds to the repeat distance of the α-helix. This periodicity will be examined more quantitatively in a later section.

The average amino acid composition may be determined for the membrane-spanning α-helixes for the RCs from *Rb. sphaeroides, Rb. capsulatus, R. rubrum,* and *Rps. viridis*. A total of 808 residues are in the membrane-spanning regions of the A, B, C, D, and E helixes of the L and M subunits of these four RCs (as aligned in Ref. 13). This analysis is restricted to residues whose Cα atoms are in the nonpolar region of the bilayer (the z values of the Cα atoms are in the range $5\ \text{Å} < z < 35\ \text{Å}$ in the convention of Figures 5 and 6). Assignment of residue location in the membrane is based on the RC structure from *Rb. sphaeroides,* and is assumed to remain valid for the other three purple bacteria. The amino acid composition, in order of decreasing abundance (in percent of total residues in the indicated region) is: Leu(15%), Ala(14%), Phe(12%), Ile(10%), Gly(9%), Val(7%), Trp(6%), Ser(6%), Thr(5%), Met(4%), Pro(3%), Arg(2%), Cys(2%), Tyr(1%), His(1%), Asn(1%), Gln(1%), Glu(1%), Asp(0%), Lys(0%). As expected, there are a large number of apolar residues, and only very few charged residues, in the membrane-spanning region of the RC.

The distribution of residues between different environments present within the membrane may also be analyzed from the sequence and structural alignments. Different amino acids exhibit different preferences between the exposed surface positions and the buried interior sites. Of the most abundant amino acids in the membrane, the apolar residues Leu, Ile, Phe, and Val tend to be located on the side of the helix exposed to the membrane, whereas Trp, Thr, and Ser, show no particular preference between the interior and surface sides. Ala and Gly prefer to be located on the helix side facing the protein interior. Surface-facing residues are defined as having >20% of their surface area exposed to the membrane in the RC structure of *Rb. sphaeroides*.

Comparison of aligned sequences from *Rb. sphaeroides, Rb. capsulatus, Rps. viridis,* and *R. rubrum* indicates that 35% (71/202) of residues in the transmembrane helixes of the L and M subunits are identical in all four sequences (10, 13). Again, this analysis considers only residues in the nonpolar region of the bilayer. Significant variation in the pattern of sequence conservation between buried and membrane-exposed residues of the transmembrane helixes (Figure 6) is observed (10). 46% (52/112) of all buried residues are identical in all four sequences, whereas only 10% (5/50) of residues with more than half of their area exposed to the membrane are conserved. This suggests that fewer restrictions are placed on residues that are

j	1 2 3	4 5 6	7 8 9	10 11 12	13 14 15	16 17 18	19 20 21
Rb. sphaeroides	SLG	VLS	LFS	GLM	WFF	TIG	IWF
Rb. capsulatus	IAG	TVS	LAF	GAA	WFF	TIG	VWY
R. rubrum	TTG	VLS	LVF	GFF	AIE	IIG	FNL
Rps. viridis	ASG	IAA	FAF	GST	AIL	IIL	FNM
Surface Exp	+	++	+	+	++	+	+
Vj	441	332	232	144	223	212	324

Figure 7 Calculation of the variability profile for the region of the MA helix in the nonpolar region of the bilayer. The sequence alignment was taken from Ref. 13. The relative residue position in this alignment is indicated by j; V_j is the variability index (number of different amino acids) for position j; and a surface exposure of "+" indicates that more than 50% of the surface area of the residue is exposed to the membrane in the RC from *Rb. sphaeroides*.

exposed to the membrane, indicating that there are few specific interactions between protein and lipid. The high tolerance to substitution of residues exposed to the membrane is analogous to the situation in globular proteins, whose surface residues also have a higher tolerance to substitutions than the buried residues (73, 74).

The periodicity of residues in a surface α-helical structure that are exposed to the membrane, coupled with the increased sequence variability of exposed residues, suggests the possibility of identifying exposed residues by analyzing the sequence alignments of homologous proteins. Assuming (*a*) that the sequence represents a transmembrane helix and (*b*) that the helix is positioned at the protein surface of the membrane-spanning region, then residues in contact with the bilayer may be identified from the pattern of hypervariable positions occurring with a periodicity of about 3.6 residues in a family of sequence alignments.

Fourier transform methods provide a quantitative approach for characterizing the periodicity of conserved and variable residues in a family of aligned sequences (13). The first step in this process is to construct a variability profile, V, for a particular family of sequences. The V_j element of this profile is defined by the number of different types of amino acid residues that are observed at a given position j in a family of aligned sequences. Construction of V for the MA helix of the RC is illustrated in Figure 7. Qualitative inspection of this profile suggests that more variable positions are associated with membrane exposed positions. To search for periodicities in V, the Fourier transform power spectra, $P(\omega)$, of V is calculated:

$$P(\omega) = \left[\sum_{j=1}^{N} (V_j - \bar{V}_j) \cos(j\omega)\right]^2 + \left[\sum_{j=1}^{N} (V_j - \bar{V}_j) \sin(j\omega)\right]^2 \quad 4.$$

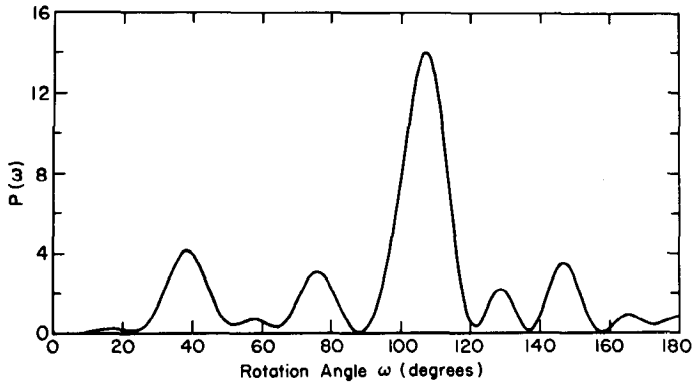

Figure 8 The Fourier transform power spectrum, P(ω) (Eq. 4), calculated for the variability profile of the transmembrane MA helix (Figure 7). The peak at ω = 105° corresponds to a periodicity of 3.4 residues/turn, and is consistent with an α-helical conformation for this sequence.

where N is the number of residues in the sequence; ω is the angular rotation angle between residues around a helical axis (it equals 100° for an ideal helix); and $\bar{V}_j$ is the mean value of V_j for the entire sequence. Similar expressions have been used by Eisenberg et al (75) and Cornette et al (76) to describe periodicity in the hydrophobicity profiles of proteins.

The P(ω) curve calculated from the variability profile for the MA helix is illustrated in Figure 8. The prominent peak near ω = 105° corresponds approximately to the periodicity of an ideal α-helix. Most of the P(ω) curves for the RC transmembrane helixes exhibit maxima at values of ω somewhat larger than the 100° expected for an ideal α-helix. In contrast, the periodicity in the hydrophobicity index of residues in helixes from soluble proteins corresponds to a value ω = 97.5° (76). The larger value of ω observed for membrane helixes may represent a slight overwinding of the helix, or more plausibly, a systematic shift in exposed residues due to interactions with adjacent helixes. For example, ω = 103° describes the periodicity of exposed residues in a coiled coil pair of α-helixes (77).

The α-helical character of the P(ω) curve may be described by the parameter ψ (13), which is defined by the average value of P(ω) in the α-helical range (90° < ω < 120°), relative to the average value of P(ω) over the entire range:

$$\psi = \left[(1/30)\int_{90°}^{120°} P(\omega)d\omega\right] \Big/ \left[(1/180)\int_{0°}^{180°} P(\omega)d\omega\right] \qquad 5.$$

Larger values of ψ correspond to a greater fraction of the P(ω) curve in the α-helical region. The following values of ψ were found for the A, B, C, D, and E helixes, respectively: 2.3, 2.9, 1.9, 1.6, and 0.9. These values of ψ were calculated using sequence alignments for the RCs from *Rb. sphaeroides, Rb. capsulatus, Rps. viridis,* and *R. rubrum,* and combining both the L and M subunit sequences. The more peripheral helixes (A and B) have larger values of ψ than the core helixes (D and E). This is consistent with the analysis that membrane-exposed residues are more poorly conserved than buried residues. These results show that ψ provides a measure of the surface exposure of a helix. This might prove useful in deriving additional information about the three-dimensional structure of membrane proteins from sequence data.

The enhanced variability of residues on the surface of an α-helix that is exposed to solvent, compared to those that face the interior side, provides an approach for identifying the topology of membrane-spanning helixes. First, the variability profile is constructed from aligned sequences of the helical regions. Next, the residue positions with greatest variability consistent with an α-helical periodicity are determined by fitting a cosine curve with $\omega = 100°$ to the variability profile. The residue positions for which this Fourier series has the greatest amplitude correspond to the most variable positions. Calculation of these positions for the 11 RC transmembrane helixes shows a strong correlation between the most variable positions and the exposed positions illustrated in Figure 6. Consequently, it is possible to assign surface-exposed sides of helixes on the basis of sequence conservation alone, without consideration of the chemical nature of the different amino acids.

The variability profile may also be used to predict the presence of α-helical segments, which are usually identified from hydropathy plots or hydrophobic moment analyses (29, 30, 35). The procedure is as follows: ψ values are calculated for a sequence contained within a window of defined size (typically 11–19 residues long). This window is moved along the sequence one residue at a time; at each position the value of ψ is determined. Regions of high ψ values (greater than ~2) correspond to a sequence that shows a strong helical periodicity and hence can be associated with a surface helix. A plot of ψ vs residue number for an 11 sequence alignment of homologous RC proteins (21) is illustrated in Figure 9. Sequence regions corresponding to the A and B helixes are evident as regions of high ψ. Local peaks in ψ are also associated with the C and D helixes, as well as the helical ab and cd segments of the interrupted helix on the periplasmic surface of the RC. Thus, this method can be used to make surface helical assignments, without any consideration of the chemical nature of the different amino acids. The requirement for surface helixes implies that this method is not applicable to α-helixes that are either completely buried inside protein or that are completely surrounded by lipid. Although we discussed only α-helixes in membrane proteins, these methods

may also be applicable to soluble proteins, and to the characterization of surface-exposed β-sheets.

An experimental approach for determining the variability profile of a sequence that does not require a number of homologous sequences has recently been described (78). Techniques of site-directed mutagenesis provide experimental procedures to determine the number of different amino acids that are tolerated at a given sequence position. Positions that accept only a small number of different amino acids are classified as having a low variability index, while positions that accept a large number of different amino acids have a high variability index. In studies on the λ repressor (a water-soluble protein), the number of substitutions that were allowed at a specific sequence position was approximately proportional to the surface exposure of that residue. This conclusion, as well as the observation that "α-helical and β-strand regions might be recognized by characteristic patterns" of high and low variability (78), is consistent with the above observations derived from sequence alignments of homologous proteins.

CONCLUDING REMARKS

A knowledge of the structures of bacterial RCs represents only the first stage in understanding the folding and properties of membrane proteins. Important questions remain concerning the actual folding and assembly mechanism for the RC in the bacterial membrane and the generality of the conclusions

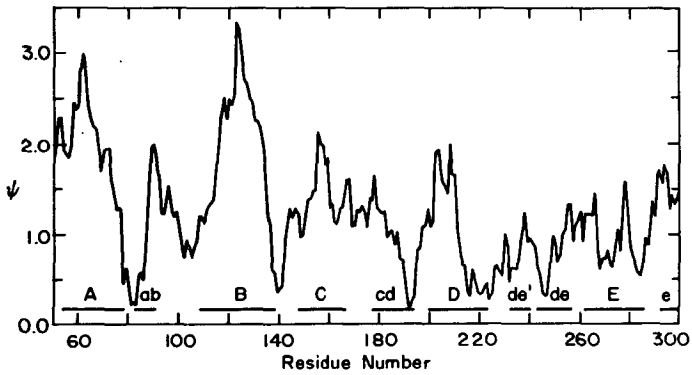

Figure 9 Calculation of ψ (Eq. 5) for a sliding window of 19 residues moving along an alignment of 11 homologous RC sequences (from Ref. 21). The sequences include both L and M subunits from bacterial RCs, and the related D1 and D2 proteins from plant photosystem II. Residue numbers correspond to the sequence of the M subunit from *Rb. sphaeroides*. Location in the sequence of α-helixes are indicated by the labeled horizontal bars. ψ is a measure of the preferential conservation of residues on one side of a (surface) helix. Large peaks for the A and B helixes are consistent with their location on the periphery of the transmembrane region of the RC.

described here for the folding and structure of other membrane proteins. The next best characterized membrane protein, bacteriorhodopsin, is believed to have a polar interior (79), in contrast to the apolar interior of the membrane-spanning region of the RCs. Whether RCs and bacteriorhodopsin represent two distinct structural motifs, or are simply limiting cases of a variety of intermediate cases, can only be decided as other high-resolution structures of membrane proteins become available. As with soluble proteins, a variety of folding patterns for membrane proteins are anticipated. The structure of the outer membrane protein porin, for which diffraction quality crystals are available, should be of special interest since this protein is believed to have a predominantly β-sheet organization (80), in contrast to the helical structures that have so far been described. Recent progress in structural and molecular biology techniques holds great promise for rapid progress in the characterization of other membrane proteins.

ACKNOWLEDGMENTS

The authors thank D. Eisenberg, L. DeAntonio, A. Chirino, J. Deisenhofer, W. DeGrado, and D. C. Wiley for discussions and comments. Research in the laboratories of the authors has been supported by grants from NIH and NSF. DCR is an A. P. Sloan research fellow.

Literature Cited

1. Michel, H. 1982. *J. Mol. Biol.* 158:567–72
2. Deisenhofer, J., Epp, O., Miki, K., Huber, R., Michel, H. 1984. *J. Mol. Biol.* 180:385–98
3. Deisenhofer, J., Epp, O., Miki, K., Huber, R., Michel, H. 1985. *Nature* 318:618–24
4. Michel, H., Epp, O., Deisenhofer, J. 1986. *EMBO J.* 5:2445–51
5. Deisenhofer, J., Michel, H. 1988. See Ref. 33, pp. 1–3
6. Allen, J. P., Feher, G. 1984. *Proc. Natl. Acad. Sci. USA* 81:4795–99
6a. Feher, G. 1983. *Meet. Biophys. Soc. (natl. lecture), Feb. 13–16, 1983, San Diego, Calif*
7. Allen, J. P., Feher, G., Yeates, T. O., Rees, D. C., Deisenhofer, J., et al. 1986. *Proc. Natl. Acad. Sci. USA* 83:8589–93
8. Allen, J. P., Feher, G., Yeates, T. O., Komiya, H., Rees, D. C. 1987. *Proc. Natl. Acad. Sci. USA* 84:5730–34
9. Allen, J. P., Feher, G., Yeates, T. O., Komiya, H., Rees, D. C. 1987. *Proc. Natl. Acad. Sci. USA* 84:6162–66
10. Yeates, T. O., Komiya, H., Rees, D. C., Allen, J. P., Feher, G. 1987. *Proc. Natl. Acad. Sci. USA* 84:6438–42
11. Yeates, T. O., Komiya, H., Chirino, A., Rees, D. C., Allen, J. P., Feher, G. 1988. *Proc. Natl. Acad. Sci. USA* 85:7993–97
12. Allen, J. P., Feher, G., Yeates, T. O., Komiya, H., Rees, D. C. 1988. *Proc. Natl. Acad. Sci. USA* 85:8487–91
13. Komiya, H., Yeates, T. O., Rees, D. C., Allen, J. P., Feher, G. 1988. *Proc. Natl. Acad. Sci. USA* 85:9012–16
14. Allen, J. P., Feher, G., Yeates, T. O., Komiya, H., Rees, D. C. 1988. See Ref. 33, pp. 5–11
15. Chang, C.-H., Schiffer, M., Tiede, D. M., Smith, U., Norris, J. 1985. *J. Mol. Biol.* 186:201–3
16. Chang, C.-H., Tiede, D., Tang, J., Smith, U., Norris, J. R., Schiffer, M. 1986. *FEBS Lett.* 205:82–86
17. Tiede, D. M., Budil, D. E., Tang, J., El-Kabbani, O., Norris, J. R., et al. 1988. See Ref. 33, pp. 13–20
18. Okamura, M. Y., Feher, G., Nelson, N. 1982. In *Photosynthesis*, ed. Govindjee, pp. 195–272. New York: Academic
19. Williams, J. C., Steiner, L. A., Ogden,

R. C., Simon, M. I., Feher, G. 1983. *Proc. Natl. Acad. Sci. USA* 80:6505–9
20. Williams, J. C., Steiner, L. A., Feher, G., Simon, M. I. 1984. *Proc. Natl. Acad. Sci. USA* 81:7303–7
21. Williams, J. C., Steiner, L. A., Feher, G. 1986. *Proteins* 1:312–25
22. Youvan, D. C., Bylina, E. J., Alberti, M., Begusch, H., Hearst, J. E. 1984. *Cell* 37:949–57
23. Michel, H., Weyer, K. A., Gruenberg, H., Lottspeich, F. 1985. *EMBO J.* 4:1667–72
24. Michel, H., Weyer, K. A., Gruenberg, H., Dunger, I., Oesterhelt, D., Lottspeich, F. 1986. *EMBO J.* 5:1149–58
25. Bélanger, G., Bérard, J., Corriveav, P., Gingras, G. 1988. *J. Biol. Chem.* 263:7632–38
26. Kauzmann, W. 1959. *Adv. Protein Chem.* 14:1–63
27. Singer, S. J. 1962. *Adv. Protein Chem.* 17:1–68
28. Henderson, R., Unwin, P. N. T. 1975. *Nature* 257:28–32
29. Eisenberg, D. 1984. *Annu. Rev. Biochem.* 53:595–623
30. Engelman, D. M., Steitz, T. A., Goldman, A. 1986. *Annu. Rev. Biophys. Biophys. Chem.* 15:321–53
31. Michel-Beyerle, M. E., ed. 1985. *Antennas and Reaction Centers of Photosynthetic Bacteria*. Berlin: Springer-Verlag
32. Biggens, J., ed. 1987. *Progress in Photosynthesis Research*. Dordrecht: Martinus Nijhoff
33. Breton, J., Vermeglio, A., eds. 1988. *The Photosynthetic Bacterial Reaction Center*. New York: Plenum
34. Kirmaier, C., Holten, D. 1987. *Photosynthesis Res.* 13:225–60
35. Kyte, J., Doolittle, R. F. 1982. *J. Mol. Biol.* 157:105–32
36. Garavito, R. M., Rosenbusch, J. P. 1980. *J. Cell Biol.* 86:327–29
37. Michel, H. 1983. *Trends Biochem. Sci.* 8:56–59
38. Garavito, R. M., Rosenbusch, J. P. 1986. *Methods Enzymol.* 125:309–28
38a. Allen, J. P., Feher, G. 1989. In *Crystallization of Membrane Proteins*, ed. H. Michel. Boca Raton: CRC. In press
39. Frank, H. A., Taremi, S. S., Knox, J. R. 1987. *J. Mol. Biol.* 198:139–41
40. Ducruix, A., Reiss-Husson, F. 1987. *J. Mol. Biol.* 193:419–21
41. Jones, T. A. 1985. *Methods Enzymol.* 115:157–71
42. Lesk, A. M., Hardman, K. D. 1982. *Science* 216:539–40
43. Bachofen, R., Wiemken, V. 1986. In *Photosynthesis III*, ed. L. A. Stahelin, C. J. Artnzen, pp. 620–31. Berlin: Springer-Verlag
44. McElroy, J. D., Feher, G., Mauzerall, D. C. 1969. *Biochim. Biophys. Acta* 172:180–83
45. Norris, J. R., Uphaus, R. A., Crespi, H. L., Katz, J. J. 1971. *Proc. Natl. Acad. Sci. USA* 68:625–28
46. Feher, G., Hoff, A. J., Isaacson, R. A., Ackerson, L. C. 1975. *Ann. NY Acad. Sci.* 244:239–59
47. Norris, J. R., Scheer, H., Katz, J. J. 1975. *Ann. NY Acad. Sci.* 244:260–80
48. Knapp, E. W., Fischer, S. F., Zinth, W., Sander, M., Kaiser, W., et al. 1985. *Proc. Natl. Acad. Sci. USA* 82:8463–67
49. Debus, R. J., Feher, G., Okamura, M. Y. 1985. *Biochemistry* 24:2488–500
50. Richards, F. M. 1977. *Annu. Rev. Biophys. Bioeng.* 6:151–76
51. Israelachvili, J. N. 1985. *Intermolecular and Surface Forces*. New York: Academic
52. Lee, B., Richards, F. M. 1971. *J. Mol. Biol.* 55:379–400
53. Richmond, T. J., Richards, F. M. 1978. *J. Mol. Biol.* 119:537–55
54. Miller, S., Lesk, A. M., Janin, J., Chothia, C. 1987. *Nature* 328:834–36
55. Mandelbrot, B. B. 1983. *The Fractal Geometry of Nature*. San Francisco: Freeman
56. Lewis, M., Rees, D. C. 1985. *Science* 230:1163–65
57. Connolly, M. 1983. *J. Appl. Crystallogr.* 16:548–58
58. Bernstein, F., Koetzle, T. F., Williams, G. J. B., Meyer, E. F., Brice, M. D., et al. 1977. *J. Mol. Biol.* 112:535–42
59. Richards, F. M. 1974. *J. Mol. Biol.* 82:1–14
60. Finney, J. L. 1975. *J. Mol. Biol.* 96:721–32
61. Rashin, A. A., Iofin, M., Honig, B. 1986. *Biochemistry* 25:3619–25
62. Hol, W. G. J. 1985. *Prog. Biophys. Mol. Biol.* 45:149–95
63. Pierson, B. K., Thornber, J. P., Seftor, R. E. B. 1983. *Biochim. Biophys. Acta* 723:322–26
64. Gibbs, J. W. 1928. *Collected Works of J. W. Gibbs*. New York: Longmans
65. Berkovitch-Yellin, Z. 1985. *J. Am. Chem. Soc.* 107:8239–53
65a. Roth, M., Lewit-Bentley, A., Michel, H., Diesenhofer, J., Huber, R. 1988. *Am. Crystallogr. Assoc. Meet., Philadelphia, Pa.* Abstr. PJ34
66. Eisenberg, D., McLachlan, A. D. 1986. *Nature* 319:199–203

67. Deisenhofer, J. 1981. *Biochemistry* 20:2361–70
68. Wilson, I. A., Skehel, J. J., Wiley, D. C. 1981. *Nature* 289:366–73
69. Wolfenden, R., Liang, Y-L. 1988. *J. Biol. Chem.* 263:8022–26
69a. Vyas, N. K., Vyas, M. N., Quiocho, F. A. 1988. *Science* 242:1290–95
70. Pape, E. H., Menke, W., Weick, D., Hosemann, R. 1974. *Biophys. J.* 14:221–32
71. Marinetti, G. V., Cattieu, K. 1981. *Chem. Phys. Lipids* 28:241–51
72. Lewis, B. A., Engelman, D. M. 1983. *J. Mol. Biol.* 166:211–17
73. Smith, E. L. 1968. *Harvey Lect.* 62:231–56
74. Chothia, C., Lesk, A. M. 1986. *EMBO J.* 5:823–26
75. Eisenberg, D., Weiss, R. M., Terwilliger, T. 1984. *Proc. Natl. Acad. Sci. USA* 81:140–44
76. Cornette, J. L., Cease, K. B., Margalit, H., Spouge, J. L., Berzofsky, J. A., DeLisi, C. 1987. *J. Mol. Biol.* 195:659–85
77. Crick, F. H. C. 1953. *Acta Crystallogr.* 6:689–97
78. Reidhaar-Olson, J. F., Sauer, R. T. 1988. *Science* 241:53–57
79. Engelman, D. M., Zaccai, G. 1980. *Proc. Natl. Acad. Sci. USA* 77:5894–98
80. Rosenbusch, J. P. 1974. *J. Biol. Chem.* 249:8019–29

Annual Review of Biochemistry
Volume 58, 1989

CONTENTS

(continued) v